Osteoporosis International
https://doi.org/10.1007/s00198-018-4686-3

ORIGINAL ARTICLE

Radiofrequency echographic multispectrometry compared with dual X-ray absorptiometry for osteoporosis diagnosis on lumbar spine and femoral neck

M. Di Paola[1,2] · D. Gatti[3] · O. Viapiana[3] · L. Cianferotti[4] · L. Cavalli[4] · C. Caffarelli[5] · F. Conversano[1] · E. Quarta[6] · P. Pisani[1] · G. Girasole[7] · A. Giusti[7] · M. Manfredini[8] · G. Arioli[8] · M. Matucci-Cerinic[9] · G. Bianchi[7] · R. Nuti[5] · S. Gonnelli[5] · M. L. Brandi[4] · M. Muratore[6] · M. Rossini[3]

Received: 9 January 2018 / Accepted: 21 August 2018

Abstract

Summary An innovative, non-ionizing technique to diagnose osteoporosis on lumbar spine and femoral neck was evaluated through a multicenter study involving 1914 women. The proposed method showed significant agreement with reference gold standard method and, therefore, a potential for early osteoporosis diagnoses and possibly improved patient management.

Introduction To assess precision (i.e., short term intra-operator precision) and diagnostic accuracy of an innovative non-ionizing technique, REMS (Radiofrequency Echographic Multi Spectrometry), in comparison with the clinical gold standard reference DXA (dual X-ray absorptiometry), through an observational multicenter clinical study.

Methods In a multicenter cross-sectional observational study, a total of 1914 postmenopausal women (51–70 years) underwent spinal ($n = 1553$) and/or femoral ($n = 1637$) DXA, according to their medical prescription, and echographic scan of the same anatomical sites performed with the REMS approach. All the medical reports (DXA and REMS) were carefully checked to identify possible errors that could have caused inaccurate measurements: erroneous REMS reports were excluded, whereas erroneous DXA reports were re-analyzed where possible and otherwise excluded before assessing REMS accuracy. REMS precision was independently assessed.

Results In the spinal group, quality assessment on medical reports produced the exclusion of 280 patients because of REMS errors and 78 patients because of DXA errors, whereas 296 DXA reports were re-analyzed and corrected. Analogously, in the femoral group there were 205 exclusions for REMS errors, 59 exclusions for DXA errors, and 217 re-analyzed DXA reports. In the resulting dataset ($n = 1195$ for spine, $n = 1373$ for femur) REMS outcome showed a good agreement with DXA: the average difference in bone mineral density (BMD, bias ± 2SD) was -0.004 ± 0.088 g/cm^2 for spine and -0.006 ± 0.076 g/cm^2 for femur. Linear regression showed also that the two methods were well correlated: standard error of the estimate (SEE) was 5.3% for spine

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00198-018-4686-3) contains supplementary material, which is available to authorized users.

✉ M. Di Paola
m.dipaola@ifc.cnr.it

[1] National Research Council, Institute of Clinical Physiology, Lecce, Italy

[2] Consiglio Nazionale delle Ricerche, Istituto di Fisiologia Clinica (CNR-IFC), Campus Ecotekne (Ed. A7), via per Monteroni, 73100 Lecce, Italy

[3] Rheumatology Unit, Department of Medicine, University of Verona, Verona, Italy

[4] Department of Surgery and Translational Medicine, University of Florence, Metabolic Bone Diseases Unit, University Hospital of Florence, Florence, Italy

[5] Department of Medicine, Surgery and Neurosciences, University of Siena, Siena, Italy

[6] O.U. of Rheumatology, "Galateo" Hospital, San Cesario di Lecce ASL-LE, Lecce, Italy

[7] SC Rheumatology, ASL 3 Genovese, Genoa, Italy

[8] Department of Neurosciences and Rehabilitation, "Carlo Poma" Hospital, ASST-Mantova, Mantova, Italy

[9] Department of Experimental and Clinical Medicine, University of Florence & SOD Rheumatology AOUC, Florence, Italy

Published online: 04 September 2018

and 5.8% for femur. REMS precision, expressed as RMS-CV, was 0.38% for spine and 0.32% for femur.
Conclusions The REMS approach can be used for non-ionizing osteoporosis diagnosis directly on lumbar spine and femoral neck with a good level of accuracy and precision. However, a more rigorous operator training is needed to limit the erroneous acquisitions and to ensure the full clinical practicability.

Keywords Diagnosis · DXA · Femoral neck · Hip · Lumbar spine · Osteoporosis · REMS · Ultrasound

Introduction

Osteoporosis is estimated to affect 200 million people worldwide [1], causing about 9 million fractures annually (i.e., an osteoporotic fracture every 3 s) [2]. The most common and disabling fractures are those of vertebrae and proximal femur [3]. Vertebral fractures are often associated with symptoms of pain, disability, and deformity [4], with postural changes associated with kyphosis that may limit activity, and lumbar fractures that may alter abdominal anatomy, leading to abdominal pain and reduced appetite [3]. Femoral fractures are associated with up to 36% excess mortality within 1 year [5], with approximately 20% of patients requiring long-term nursing home care and only 40% fully regaining their pre-fracture level of independence [6].

Prevention, detection and treatment of osteoporosis should therefore be a mandate of primary care providers. Nevertheless, it has become increasingly clear that many patients are not being given appropriate information about prevention and many patients are not receiving appropriate testing to diagnose osteoporosis or establish osteoporosis risk [7]. Osteoporosis is actually preventable and treatable, but because there are no warning signs prior to a fracture, many people are not being diagnosed in time to receive effective therapy and/or the appropriate lifestyle corrections during the early phase of the disease. An International Osteoporosis Foundation (IOF) survey, conducted in 11 countries, showed that restricted access to diagnosis before the first fracture is one of the main causes of osteoporosis underdiagnosis and undertreatment [8].

The diagnosis of osteoporosis is currently established by measurements of bone mineral density (BMD) or by the occurrence of adulthood femoral or vertebral fracture in the absence of major trauma. Dual X-ray absorptiometry (DXA) of proximal femur and lumbar spine is the reference technology used to establish or confirm a diagnosis of osteoporosis.

However, at least in some European countries, the limited accessibility of DXA examinations, mainly due to limited number of densitometers, restrictions in personnel permitted to perform scans, and/or absence of reimbursement, has been reported as an important problem limiting the healthcare system effectiveness in osteoporosis diagnosis and fracture prevention [9].

For these reasons, several quantitative ultrasound (QUS) methods for non-ionizing osteoporosis diagnosis have been developed and investigated in several studies [10–24]. Nevertheless, most of these approaches are applicable only to peripheral skeletal sites (e.g., calcaneus, tibia, radius, etc.), thus, providing only indirect estimations of the actual bone health status at the axial reference sites for osteoporosis diagnosis (i.e., lumbar vertebrae and femoral neck). On the other hand, the actual clinical translation of the few methods working on the axial sites was prevented either by the limited accuracy [18, 19] or by bulkiness and complexity of the employed device [20], which actually did not effectively overcome most of DXA limitations.

An innovative echographic approach for osteoporosis diagnosis, directly applicable on both femoral neck and lumbar spine, has been recently introduced and clinically validated through single-center studies [25, 26]. This developed approach has been subsequently defined as Radiofrequency Echographic Multi Spectrometry (REMS). The main output parameter of this fully non-ionizing technique is BMD_{US}, a diagnostic index expressed as grams/cm^2, which is measured directly on lumbar vertebrae or proximal femur and has shown significant correlations with the corresponding BMD values and good agreement levels with DXA-based diagnoses assumed as the gold standard reference [25–27].

However, recent literature [28] has pointed out, at least in some cases, a suboptimal adherence of clinical routine DXA scans to the International Society for Clinical Densitometry (ISCD) guidelines. Although the magnitude of the bias introduced by such errors has not been investigated, a careful check by experienced operators has been advisable to assume clinical routine DXA reports as a fully reliable reference gold standard. Furthermore, in order to exploit the maximum power of the diagnostic capabilities of the REMS technology, it is important to fully comply to the echographic scan procedure with the protocol and the indications provided by the device manufacturer.

The aim of this multicenter cross-sectional observational study was to assess "short-term intra-operator precision" (from here on identified as "precision"), "short-term inter-operator repeatability" (from here on identified as "repeatability") and diagnostic accuracy of the REMS investigations in comparison with the clinical gold standard reference DXA. Specific attention has been paid to verify the actual compliance of the performed DXA and REMS scans with the corresponding guidelines for the correct clinical execution and to assess the effect of possible errors on the diagnostic performance of the innovative technique under evaluation.

Methods

Patients

The study was a multicenter cross-sectional observational study on postmenopausal women fulfilling the following enrollment criteria: Caucasian ethnicity, aged 51–70 y, body mass index (BMI) < 40 kg/m^2, absence of significant deambulation impairment (e.g., patients on wheelchairs were excluded in order to avoid lengthening of the examination procedure and slowdown in the hospital workflow), and medical prescription for a spinal and/or femoral DXA investigation.

The patients were recruited from six Italian centers: "Galateo" Hospital (San Cesario di Lecce, Lecce), "La Colletta" Hospital (Arenzano, Genoa), "Le Scotte" University Hospital (Siena), "Careggi" University Hospital (Florence), "Villa Monna Tessa" University Hospital (Florence), and "Borgo Roma Gianbattista Rossi" University Hospital (Verona).

Inclusion started in November 2016 and ended in September 2017, recruiting a total of 1914 patients. The enrolled patients underwent a spinal and/or femoral DXA investigation (according to their medical prescription) and an echographic scan of the same anatomical sites performed with the REMS approach, as detailed in the following paragraphs.

The study protocol was approved by the Ethics Review Boards of all the participating hospitals. All the enrolled patients voluntarily entered the study after giving written informed consent.

DXA measurements

Anteroposterior DXA scans were performed according to the standard clinical routine procedures employing one of the following devices: Discovery W (Hologic, Waltham, MA, USA), Delphi A (Hologic), Horizon A (Hologic), QDR 4500A (Hologic), or Lunar Prodigy (GE Healthcare, Madison, WI, USA).

All the Hologic scanners used the following reference ranges: lumbar spine was considered "osteoporotic" if BMD ≤ 0.777 g/cm^2 (T-score ≤ −2.5), "normal" if BMD ≥ 0.932 g/cm^2 (T-score ≥ −1.0), and "osteopenic" for intermediate BMD values (−2.5 < T-score < −1.0); femoral neck was considered "osteoporotic" if BMD ≤ 0.577 g/cm^2 (T-score ≤ −2.5), "normal" if BMD ≥ 0.733 g/cm^2 (T-score ≥ −1.0), and "osteopenic" for intermediate BMD values (−2.5 < T-score < −1.0). On the other hand, the Lunar Prodigy scanner used different BMD thresholds for the same diagnostic classifications: lumbar spine was considered "osteoporotic" if BMD ≤ 0.885 g/cm^2 (T-score ≤ −2.5), "normal" if BMD ≥ 1.054 g/cm^2 (T-score ≥ −1.0), and "osteopenic" for intermediate BMD values (−2.5 < T-score < −1.0); femoral neck was considered "osteoporotic" if BMD ≤ 0.685 g/cm^2 (T-score ≤ −2.5), "normal" if BMD ≥ 0.854 g/cm^2 (T-score ≥ −1.0), and "osteopenic" for intermediate BMD values (−2.5 < T-score < −1.0). For all the employed devices, the reference curves adopted by the DXA scanner software to calculate the T-score values were integrated in the software itself and were automatically selected based on patient characteristics (always Caucasian females in the present study) and scanned anatomical site. For each considered combination of DXA scanner and anatomical site, the reference database, which could be NHANES III or a proprietary manufacturer database, calculated the T-score value employing the above-listed reference ranges. For the purposes of the present study, diagnostic classifications were always based on T-score values and, just in order to compare BMD with BMD$_{US}$, BMD values measured by the Lunar Prodigy were preliminarily converted in Hologic-equivalent values as described in the "Data analysis" section.

Spinal investigations were carried out with hip and knee both at 90° of flexion, whereas for femoral examinations the patient's femur was straight on the table, with the shaft being parallel to the vertical edge of the obtained image, and with a 15–25° internal rotation, achieved by using a dedicated positioning device.

DXA medical reports always included both the BMD value of the considered anatomical site, expressed as grams per square centimeter (g/cm^2), and the corresponding T-score value, based on the standard reference database for Caucasian women integrated in the specific DXA scanner software.

All the DXA medical reports were anonymized and digitally stored for the subsequent analyses.

Employed DXA scanners underwent daily quality control and regular maintenance for the whole study period.

REMS acquisitions

REMS scans of lumbar vertebrae and proximal femur were performed employing a dedicated echographic device (EchoStation, Echolight Spa, Lecce, Italy), equipped with a convex transducer operating at the nominal frequency of 3.5 MHz and used as recommended by the manufacturer. Data processing methodologies implemented in the REMS approach are those detailed in previous papers [25, 26]. In particular, the EchoStation software (EchoStudio, Echolight Spa, Lecce, Italy) integrates a proprietary database of reference ultrasound spectral models, which was built as described in the cited papers [25, 26] and is used to calculate BMD$_{US}$ values, and a normative reference database (National Health and Nutrition Examination Survey, NHANES), which is used to derive the corresponding T-score and Z-score values. In short, the adopted approach is based on the calculation of the Osteoporosis Score [25, 26], which is derived starting from the insonification of the target bone volume and corresponds to the percentage of analyzed bone regions that were classified as "osteoporotic" through the dedicated spectral analyses. Linear equations are then employed to transform

Osteoporosis Score into BMD_{US} values, which are finally expressed also as T-score and Z-score values through quantitative comparisons with the NHANES reference curves.

All the employed echographic devices, which were brand new and had undergone exactly the same calibration procedure, were provided in a modified research configuration, which, for each completed acquisition, automatically stored not only the final medical report, but also the corresponding sequence of B-mode images and the related unprocessed "raw" ultrasound signals.

Lumbar scans were performed by placing the echographic transducer in a trans-abdominal position under the sternum (Fig. 1a), in order to initially visualize L1 lumbar vertebra and then moving it until L4 according to the on-screen and audible indications provided by the device software (EchoStudio, Echolight Spa, Lecce, Italy). Each lumbar scan lasted 80 s (20 s per vertebra) and it was followed by an automatic processing time of about 1–2 min. Figure 1b shows a typical echographic image that can be seen during a clinical acquisition: there are actually more vertebral interfaces visible, but each of them will be acquired during a dedicated 20-second step of the scan procedure during which the vertebra has to be located in the central part of the B-mode image.

Proximal femur scans were performed by placing the echographic transducer parallel to head-neck axis of the femur, in order to visualize the typical proximal femur profile, including the interfaces of femoral head, neck, and trochanter (Fig. 2). Once found the proper visualization, and started the acquisition, the operator had just to hold this image for 40 s, according to the on-screen and audible indications provided by the EchoStudio software, and then wait for about 1 min for the automatic data processing.

For all the performed vertebral and femoral acquisitions, transducer focus (21–100 mm), and scan depth (60–210 mm) were adjusted for each patient in order to have the target bone interface (i.e., vertebral surface or femoral neck) in the ultrasound beam focal zone and at about halfway through the image depth.

In each clinical center participating in the study, for both the considered anatomical sites, the first 10 enrolled patients underwent two consecutive REMS scans, with patient repositioning between the scans, in order to assess the measurement precision and repeatability. In particular, the first 5 patients underwent two consecutive investigations performed by an experienced operator and the corresponding data were included in the precision assessment, whereas the subsequent 5 patients underwent two consecutive investigations performed by two different operators, an experienced one and another who had previously received only a 3-hour specific training session; these data were used to assess the repeatability. Finally, all the subsequent patients underwent a single scan per anatomical site, performed by an operator who had initially received only a 3-hour specific training, and the corresponding data were used only for diagnostic accuracy evaluations.

All the REMS medical reports, together with the corresponding echographic images and related raw signals, were anonymized before starting the subsequent analyses.

Data analysis

Precision For each considered anatomical site, precision was assessed as defined by Engelke and Gluer [29], using the data acquired on the first 5 patients enrolled in each of the 6 clinical centers involved in the study. Therefore, a total of 30 cases were used for calculating precision and for these calculations, in order to quantify just the maximum achievable precision, only repeated measurements performed by an experienced operator were employed, as described in the previous paragraph.

For each considered patient, standard deviation (SD) of repeated BMD_{US} measurements was calculated, and REMS precision was expressed as the root-mean-square coefficient of variation (RMS-CV). Least significant change (LSC) for a 95% confidence level was also calculated as recommended by the International Society for Clinical Densitometry (ISCD). These calculations were performed via the ISCD precision calculator (available at http://www.iscd.org/resources/calculators/).

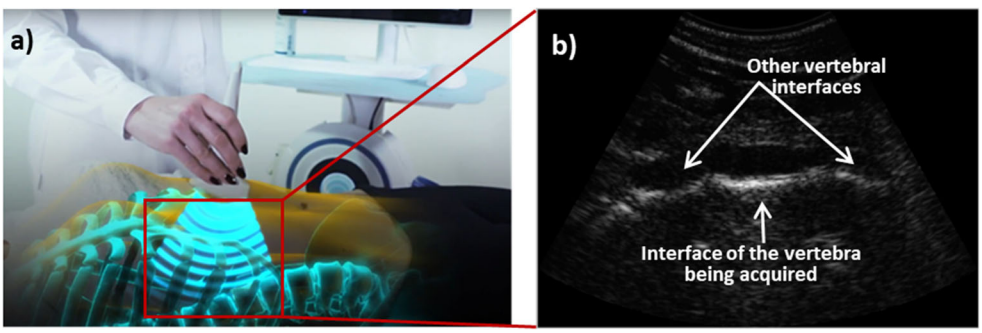

Fig. 1 Echographic transducer placement for REMS acquisitions on lumbar vertebrae: **a** schematic depiction of the trans-abdominal transducer placement and **b** visualization of a typical echographic image, with the identification of vertebral interfaces, distinguishing between the vertebra that is undergoing the acquisition and the neighboring ones

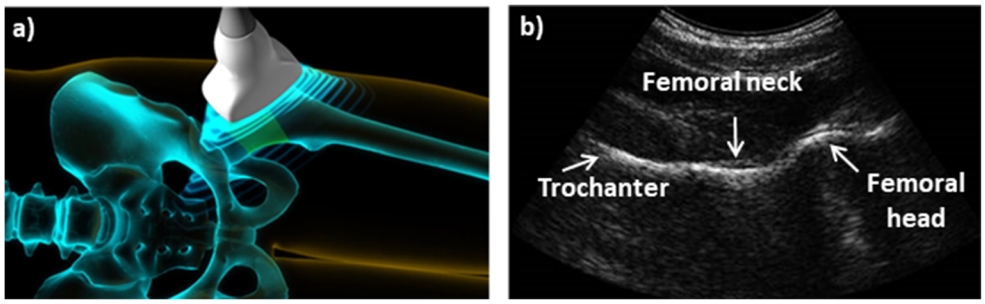

Fig. 2 Echographic transducer placement for REMS acquisitions on femoral neck: **a** schematic depiction of the transducer placement parallel to the head-neck axis of the femur and **b** visualization of a typical echographic image, with the identification of the interfaces of femoral head, neck, and trochanter

Repeatability For each considered anatomical site, repeatability was assessed on the data acquired on the second 5 patients enrolled in each of the 6 clinical centers involved in the study. Therefore, also for repeatability, a total of 30 cases were used for the calculations. In this case, in order to maximize the difference between the operators, for each of the 30 patients, we included only repeated measurements performed one by an experienced operator and the other by a newly trained one, as previously described.

Calculations were carried out similarly to those for precision (i.e., the same ISCD calculator was used, including for each considered patient two measurements performed by two different operators in the place of the two measurements performed by the same operator), and also the repeatability was expressed in terms of RMS-CV and LSC for a 95% confidence interval.

Diagnostic accuracy For the purposes of the present study, spinal DXA reports were processed separately from femoral ones. According to spinal DXA reports, each patient was classified as "osteoporotic" if lumbar T-score ≤ -2.5 and as "non-osteoporotic" if lumbar T-score > -2.5. An independent classification employing the same threshold was adopted on the basis of femoral neck T-score values obtained from femoral DXA reports. In both cases the "non-osteoporotic" patients were further classified as "osteopenic" if $-2.5 <$ T-score < -1.0 or "healthy" if T-score ≥ -1.0.

The whole classification process was independently repeated on the basis of the corresponding lumbar and femoral neck T-score values obtained from REMS scans.

In order to assure the maximum reliability of the diagnostic outputs, all the collected medical reports (DXA and REMS), with the corresponding echographic images and raw data in the REMS case, were independently checked by two experienced operators in order to identify the possible errors that could have provided improper measurements, potentially resulting in inappropriate diagnostic classifications. DXA errors were identified according to the most updated ISCD guidelines [30] and to the indications coming from recent literature [28]: they were typically associated with inaccurate patient positioning, wrong data analysis (e.g., incorrect placement of analysis boxes in the image), presence of artifacts, or mistakes in the input of demographic characteristics. REMS errors were identified as deviations from the acquisition procedure described in the EchoStation user manual: they were typically associated with wrong or suboptimal settings of transducer focus and/or scan depth, or with incomplete adherence to the on-screen and audible indications provided by the software (e.g., missing or delayed movement from a given vertebra to the subsequent one).

Both the operators, which were blind with respect to each other findings, were asked to carefully check each medical report for the possible presence of any of the previously listed error types. For instance, referring to wrong data analyses associated with DXA scans, a typical error was represented by a slight misplacement (1–2 mm) of an intervertebral line or by the inclusion in the analysis of lumbar vertebrae not belonging to L1–L4. On the other hand, referring to REMS acquisitions, taking into account that transducer focus could be set only at fixed values (e.g., 21 mm, 36 mm, 45 mm, 53 mm, etc.), a typical error was the selection of a transducer focus different from the "ideal" value. Once both the operators independently completed the report analysis, they discussed together all the cases that had received different classifications (i.e., presence/absence of errors and/or type of error) until a consensus was reached.

For each considered anatomical site, DXA reports containing errors related to wrong data analysis, presence of removable artifacts, or mistakes in the demographic inputs were re-analyzed and the corrected DXA reports were considered for the purposes of the present study, whereas the cases presenting errors due to inaccurate patient positioning or non-removable artifacts were excluded from subsequent analyses.

On the other hand, since REMS reports in principle cannot be re-analyzed and corrected offline, a third experienced operator was asked to check the REMS reports initially identified as erroneous and to verify if some errors could be neglected because they were not supposed to have a significant effect on the measurement outcome. In the end, the REMS reports identified as erroneous also by the third operator were actually excluded from the study, whereas the others were retained.

Diagnostic accuracy of the REMS approach was then assessed on the remaining patients by assuming DXA outputs as the gold standard reference and by determining sensitivity and specificity in the discrimination between "osteoporotic" and "non-osteoporotic" patients. The diagnostic concordance between the two methods was also assessed, by calculating the percentage of patients being classified in the same diagnostic category (osteoporotic, osteopenic, or healthy) by both DXA and REMS and the corresponding Cohen's kappa (k). Furthermore, the degree of correlation between DXA and REMS T-score values was quantified through a linear regression analysis, by calculating the slope of the regression line, the Pearson's correlation coefficient (r) and the coefficient of determination (r^2). Finally, we directly assessed the agreement between BMD and BMD_{US} values by measuring the standard error of the estimate (SEE) and through the Bland-Altman method [31].

In order to compare BMD with BMD_{US} values, we had to take into account that there are systematic differences in how BMD values are measured and reported among DXA scanners from various manufacturers. Since BMD_{US} already showed a very good correlation with BMD measurements performed with Hologic densitometers [25, 26], BMD values measured by Lunar scanners were preliminarily converted in Hologic-equivalent values by applying specific conversion formulas derived from literature-available papers for both lumbar spine [32] and femoral neck [33].

Results

Study population and quality assessment on medical reports

A total of 1914 postmenopausal women aged between 51 and 70 years were included in this study: according to their medical prescription, 1276 underwent both lumbar and femoral investigations, 277 were examined only on the lumbar site, and 361 only on the femoral one. Therefore, the lumbar spine study actually included 1553 patients (lumbar group), whereas 1637 patients were considered for the femoral neck study (femoral group). Based on BMI values, 55.5% of the patients in the lumbar group were normal- or under-weight, 34.5% were overweight, and 10.0% were obese; in the femoral group we had 54.6% normal- or under-weight, 34.2% overweight, and 11.2% obese.

Table 1 summarizes the average patient characteristics for each considered anatomical site, together with the results of the quality assessment on medical reports, which resulted in the exclusion of: (i) 358 patients from the lumbar group, 78 (5.0%) because of DXA errors due to inaccurate patient positioning and 280 (18.0%) because of non-recoverable REMS errors; (ii) 264 patients from the femoral group, 59 (3.6%) because of DXA errors (51 cases of inaccurate patient positioning and 8 cases with artifacts that could not be

Table 1 Average characteristics of the enrolled patients for each considered anatomical site and results of the quality assessment on medical reports

	Lumbar spine	Femoral neck
Enrolled patients	1553	1637
Age (y)	60.7 ± 5.4	60.9 ± 5.5
Height (cm)	159.4 ± 6.1	159.4 ± 6.2
Weight (kg)	63.2 ± 10.1	63.6 ± 10.5
BMI (kg/m^2)	24.9 ± 3.8	25.0 ± 4.0
Initially erroneous DXA reports	374 (24.1%)	276 (16.9%)
Re-analyzed DXA reports	296 (19.1%)	217 (13.3%)
- Wrong data analysis	210 (13.5%)	215 (13.1%)
- Correctable artifacts	84 (5.4%)	–
- Data input mistakes	2 (0.1%)	2 (0.1%)
Excluded DXA reports	78 (5.0%)	59 (3.6%)
- inaccurate patient positioning	78 (5.0%)	51 (3.1%)
- uncorrectable artifacts	–	8 (0.5%)
REMS reports initially identified as erroneous	340 (21.9%)	239 (14.6%)
Recovered REMS reports	60 (3.9%)	34 (2.1%)
- Acceptable focus selection	42 (2.7%)	27 (1.7%)
- Acceptable scan depth selection	18 (1.2%)	7 (0.4%)
Excluded REMS reports	280 (18.0%)	205 (12.5%)
- Wrong focus selection	185 (11.9%)	165 (10.1%)
- Wrong scan depth selection	92 (5.9%)	40 (2.4%)
- No adherence to scan procedure	3 (0.2%)	–

removed through re-analysis) and 205 (12.5%) because of REMS errors.

On the other hand, the re-analysis of DXA reports containing errors related to wrong data analysis, presence of removable artifacts, or mistakes in the demographic inputs, allowed the correction of 296 reports in the lumbar group and 217 reports in the femoral group.

The initial analysis of REMS reports had identified 340 errors in the lumbar group and 239 errors in the femoral group. The final check performed by the third experienced operator identified as negligible 60 errors in the lumbar group and 34 errors in the femoral group, determining the recovery of the corresponding reports.

As a final result, the REMS diagnostic accuracy was assessed on 1195 patients for lumbar spine and 1373 patients for femoral neck. The percentages of obese, overweight and normal- or under-weight patients stayed substantially unchanged for both the lumbar group and the femoral group.

The described dataset was identified as the "primary dataset" and represented the "best case" (i.e., the case in which all the REMS and DXA errors were carefully identified and the resulting REMS performance was not affected by any operator-dependent inaccuracy). These results are illustrated and discussed in the next paragraphs.

On the other hand, we also identified a "supplementary dataset" in which all the original REMS reports were included, without any error exclusion, and the corresponding diagnostic accuracy was again assessed with respect to the "cleaned" DXA data, in order to evaluate the REMS performance in a "real life" context against a gold standard. This second dataset, whose results are summarized in the "Discussion" section and presented in detail in the Supplementary Material, included 1475 patients from the lumbar group and 1578 patients from the femoral group.

Diagnostic accuracy of the REMS approach and agreement with DXA

The REMS approach applied to the "primary dataset" effectively discriminated the osteoporotic patients from the non-osteoporotic ones on both lumbar spine (sensitivity = 91.7%, specificity = 92.0%) and femoral neck (sensitivity = 91.5%, specificity = 91.8%).

The good diagnostic performance was confirmed also when all the three possible classification categories (osteoporotic, osteopenic, healthy) were considered: the diagnostic concordance between DXA and REMS was 88.8% (k = 0.824, $p < 0.001$) for lumbar spine and 88.2% ($k = 0.794$, $p < 0.001$) for femoral neck, respectively.

These results are further emphasized by the high degree of correlation between the T-score values provided by the two techniques for both lumbar spine ($r = 0.94$, $p < 0.001$) and femoral neck ($r = 0.93$, $p < 0.001$), and also by the slope of the corresponding regression lines (0.95 for spine and 0.97 for femur), as shown in Fig. 3a and b, respectively.

Figure 4 shows the Bland-Altman plots obtained to assess the differences between DXA-measured BMD values and REMS-measured BMD_{US} values for each considered anatomical site: the average difference (expressed as bias ± 2 SDs) was -0.004 ± 0.088 g/cm^2 for lumbar spine (Fig. 4a) and -0.006 ± 0.076 g/cm^2 for femoral neck (Fig. 4b). Figure 4a, b also emphasizes the absence of any visible trend linking the difference between BMD_{US} and BMD to their average value, which means that the accuracy of BMD_{US} in the estimation of BMD does not depend on the BMD value. These results, combined with the values of the coefficient of determination ($r^2 = 0.89$ for lumbar spine and $r^2 = 0.87$ for femoral neck) and with the corresponding standard errors of the estimate (SEE = 0.044 g/cm^2 [5.3%] for lumbar spine and SEE = 0.038 g/cm^2 [5.8%] for femoral neck) documented the actual strength of the relationship between BMD_{US} and BMD. Table 2 contains the full comparison between the results obtained for each of the considered anatomical sites.

All the results obtained considering the "supplementary dataset" are available as Supplementary Material.

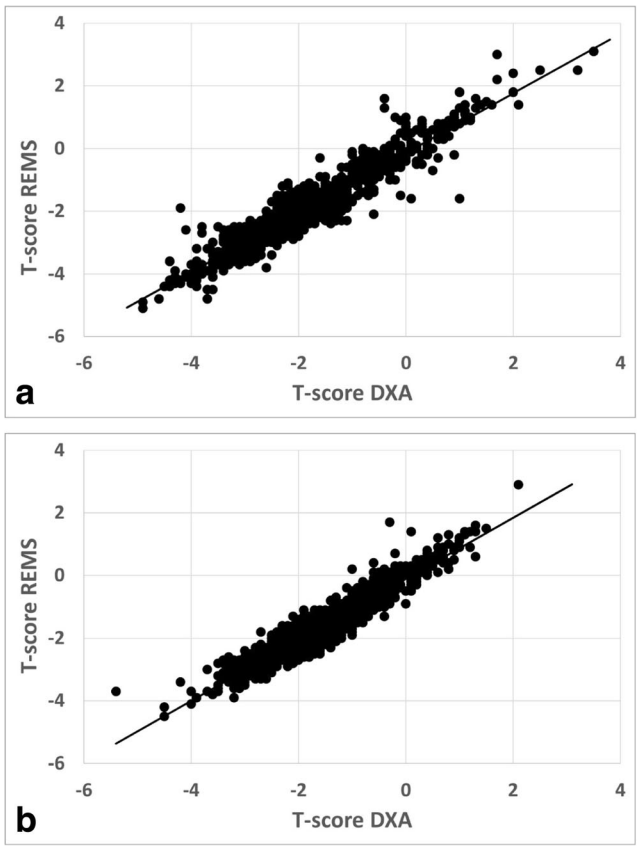

Fig. 3 Scatterplot of DXA T-score and REMS T-score for patients in the "primary dataset": **a** lumbar spine (slope of the regression line = 0.95, r = 0.94, $p < 0.001$) and **b** femoral neck (slope of the regression line = 0.97, $r = 0.93$, $p < 0.001$)

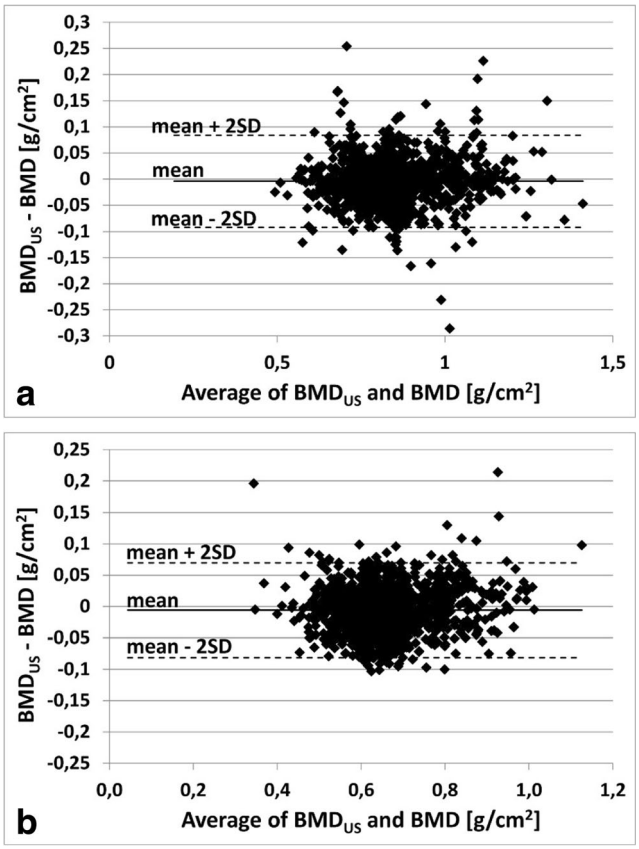

Fig. 4 Bland-Altman plot for comparison of BMD_{US} and BMD measurements for patients in the "primary dataset": **a** lumbar spine and **b** femoral neck

REMS precision and repeatability

Evaluations of precision and repeatability of REMS outcomes were also carried out, with the aim of assessing the intrinsic precision and reproducibility of the proposed method, independently of operator experience.

The patients used for the precision assessment had the following characteristics: for lumbar spine, BMI in the range 18.7–38.1 kg/m^2 (mean ± SD = 25.0 ± 3.9 kg/m^2) and age in the range 51–70 y (mean ± SD = 60.8 ± 5.2 y); for femoral neck, BMI in 18.6–34.3 kg/m^2 (24.3 ± 4.0 kg/m^2) and age in 51–69 y (58.7 ± 5.0 y). Analogously, the patients used for the repeatability assessment had the following characteristics: for lumbar spine, BMI in 19.4–33.3 kg/m^2 (24.0 ± 3.3 kg/m^2) and age in 55–69 y (61.5 ± 4.0 y); for femoral neck, BMI in 19.8–29.7 kg/m^2 (24.1 ± 2.7 kg/m^2) and age in 55–66 y (59.4 ± 3.7 y).

Precision, expressed as RMS-CV, was 0.38% (95% confidence interval: 0.28–0.48%) for lumbar spine and 0.32% (0.24–0.40%) for femoral neck, and the corresponding LSC value for a 95% confidence level was 1.05% for lumbar spine and 0.88% for femoral neck, respectively.

Analogous calculations were performed to assess repeatability, producing the following results: RMS-CV = 0.54% (0.40–0.68%) and LSC = 1.50% for lumbar spine; RMS-CV = 0.48% (0.36–0.60%) and LSC = 1.33% for femoral neck.

In order to provide a further visual insight in the variability of REMS measurements, Fig. 5 shows the box plots of the

Table 2 Results of the REMS accuracy evaluations for each considered anatomical site. In all the calculations, DXA results obtained for the retained patients after the exclusion of medical reports containing uncorrectable errors, and the re-analysis of those containing recoverable errors, were assumed as the reference ground truth

Anatomical site	Lumbar spine	Femoral neck
Retained cases (n)	1195	1373
Sensitivity	91.7%	91.5%
Specificity	92.0%	91.8%
Diagnostic concordance	88.8%	88.2%
k	0.824*	0.794*
r	0.94*	0.93*
r^2	0.89*	0.87*
Regression line slope	0.95	0.97
SEE (g/cm^2)	0.044(5.3%)	0.038(5.8%)
Average difference(bias ± 2SD, g/cm^2)	−0.004 ± 0.088	−0.006 ± 0.076

*$p < 0.001$

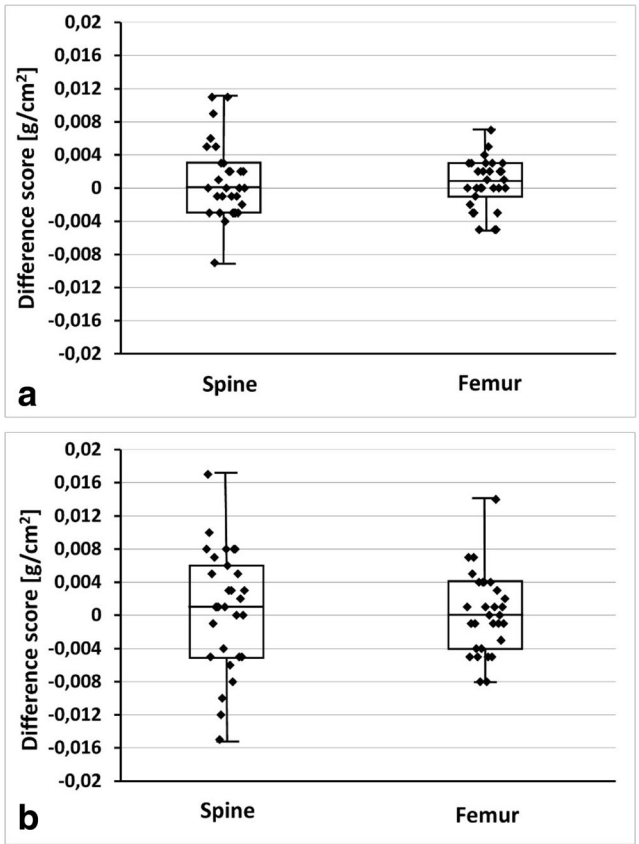

Fig. 5 Box plots of the repeated BMD_{US} measurements used for the calculation of precision and repeatability parameters: **a** intra-operator variability and **b** inter-operator variability

repeated measurement data used for the calculation of precision and repeatability parameters.

Discussion

This study assessed the diagnostic performance of REMS investigations in comparison with DXA in a multicenter clinical context, paying specific attention to avoid the possible biases due to errors that could affect the outcome of either of the two techniques. For this purpose, we considered two different datasets: (i) a "primary dataset," which was obtained by excluding all the cases containing DXA and REMS errors and represented the "best case" (i.e., the REMS diagnostic potential when a strict quality check is applied); (ii) a "supplementary dataset," which still excluded the DXA errors but included all the REMS reports excluded by the "primary dataset" and represented the "real life" case (i.e., the minimum REMS performance that can be anyway achieved even without any quality control nor patient exclusion).

In this way, we documented that, if all the guidelines and recommendations are scrupulously followed, the actual diagnostic capabilities of the REMS approach result in both sensitivity and specificity above 90% for each considered anatomical site. The latter is a particularly relevant result, since both the ISCD [34] and the UK National Osteoporosis Society (NOS) [35] recommend, for ultrasound devices to be used as pre-screening tools, the identification of two specific thresholds of ultrasound parameters such that the screening protocols identify those with and those without osteoporosis with 90% sensitivity and 90% specificity, whereas DXA scans have to be performed in addition to the ultrasound examinations only for those with ultrasound parameter values between the lower and the upper threshold [34, 35]. Therefore, the fact that the REMS approach directly provides both sensitivity and specificity above 90% by employing a unique threshold, which is the same employed by DXA (i.e., T-score = −2.5), suggests that the proposed echographic method could be used to classify all the patients into diagnostic categories, without requiring additional DXA scans. If we consider the "supplementary dataset", sensitivity decreases to 81.0% for spine and 81.7% for femur, whereas the corresponding specificities are 84.3 and 89.7%, respectively.

Another important aspect can be pointed out by analyzing the diagnostic concordance between DXA and REMS when all the three possible classification categories (osteoporotic, osteopenic, healthy) were considered, resulting in a diagnostic concordance of 88.8% ($k = 0.824$, $p < 0.001$) for lumbar spine and 88.2% ($k = 0.794$, $p < 0.001$) for femoral neck (considering the "primary dataset"). In this context, it has to be emphasized that many of the reported cases of "wrong classification" are actually associated with "borderline situations" (e.g., one patient receiving T-score = −2.5 from DXA and T-score = −2.4 from REMS was considered as an erroneous diagnosis, although the REMS result was very close to the gold standard value). To give a quantitative measure of the impact of these cases on the overall results, we can say that, if we accept a "tolerance" up to 0.3 on the REMS T-score value of "borderline" patients before labelling them as wrong classifications, the diagnostic concordance would reach 97.4% for lumbar spine and 98.0% for femoral neck. When the "supplementary dataset" was considered, the initial diagnostic concordance resulted 76.4% ($k = 0.629$, $p < 0.001$) for lumbar spine and 81.9% ($k = 0.691$, $p < 0.001$) for femoral neck, which increased respectively to 86.8 and 92.0% by accepting the 0.3 tolerance on T-score value of "borderline" patients.

Furthermore, for both the lumbar spine and femoral neck sites, the precision parameters determined in the present study for the REMS approach were better than the corresponding values typically reported in literature for the employed comparative gold standard. For instance, referring to lumbar spine, the precision RMS-CV was 0.38% for REMS, whereas it has been recently reported to be in the range 1.07–1.34% for DXA when employed on women with characteristics similar to the present study patients [36]. Analogously, for the femoral neck, a typical reported value of precision RMS-CV for DXA is 1.47% [37], whereas REMS resulted in RMS-CV = 0.32%.

On the other hand, to the best of our knowledge, there are no reported studies actually quantifying the repeatability of BMD measurements obtained through anteroposterior DXA measurements on living subjects. Among the literature-available papers, those closest to the mentioned topic are the following three: Trevisan et al. [38] evaluated the inter-device variability, reporting a CV higher than 2% for BMD measurements on volunteers; Larnach et al. [39] assessed the inter-operator repeatability of lateral spinal scans, reporting CV = 3.8% for BMD measurements; Raffan et al. [40] evaluated the inter-operator variability of body composition measurements performed on canine cadavers, reporting CVs variable in a wide range (from 0.04 to 1.6%) as a function of the considered tissue.

Regarding the capability of BMD_{US} to estimate the exact BMD value provided by DXA, one of the parameters that best emphasize the accuracy of REMS technology is probably the value of SEE expressed as a percentage, which resulted equal to 5.3% for lumbar spine and to 5.8% for femoral neck. This indicates that the average distance between the measurement output and the regression line is less than 6% of the value for both the considered anatomical sites. On the other hand, the SEE values increased to 10.0% for spine and to 9.6% for femur, respectively, when the "supplementary dataset" was considered. The fact that the worsening was more marked for spine than for femur reflected the higher percentage of erroneous REMS scans that was observed in the lumbar group. Therefore, even though REMS lumbar investigations resulted potentially more accurate than femoral ones in the "primary dataset", they also resulted somewhat more difficult

to be correctly carried out, turning out in their slightly lower accuracy when REMS errors were not excluded.

Actually, after the described re-analyses of erroneous reports, the final exclusion rates for DXA were 5.0% for spine and 3.6% for femur. On the other hand, in the case of REMS we did not have the possibility of correcting the reports through a re-analysis of the acquired data and therefore, in order to avoid obtaining a REMS performance biased by operator-dependent errors, in the "primary dataset," we had to exclude 18.0% of the patients from the lumbar group and 12.5% of the patients from the femoral group. This represents a limitation of this technique, as discussed in more detail later. A "post hoc" analysis of the average temporal distribution of the errors showed that DXA errors were almost uniformly distributed along the study period, whereas REMS errors were concentrated in the first 3–4 months, indicating the possibility of improving the clinical practicability of REMS through a more rigorous training of the operators, aimed at reducing the time span of the learning curves and the related initial error rates.

However, the results obtained from the "primary dataset" were not influenced by the temporal distribution of the errors, since all of them (for both REMS and DXA acquisitions) were either corrected or excluded. Therefore, the "primary dataset" represented the "best case" (i.e., the maximum achievable REMS diagnostic performance when no errors are made by the operator). The actual effect of operator-dependent REMS errors on the overall performance of this technique was assessed by considering also a "supplementary dataset", including all the available REMS reports and representing a "real life" context. Main results obtained through the "supplementary dataset" have been discussed in this section and further details are available as Supplementary Material.

The main limitation of the present study is actually represented by the number of REMS reports excluded because of operator-dependent errors observed in the wrong selection of depth/focus at echographic scanning. Although their incidence on REMS diagnostic performance was quantified by considering two different datasets differing from each other only for the inclusion/exclusion of the erroneous REMS reports, this could in principle question the clinical practicability of the technique for the future. However, in a perspective view, the problem will be essentially solved through the mentioned more rigorous training of the operators (a 3-day training program, for instance, will be definitely more effective than the 3-hour training adopted in the present study). Furthermore, given the intrinsic characteristics of REMS technique and settings, an additional software feature can in principle be implemented to provide an automatic real-time indication on the most suitable values of transducer focus and scan depth for the patient undergoing the examination. Nevertheless, the fact that re-analysis of REMS reports was not currently possible represents a disadvantage of this technique, also in comparison with DXA, even if the scanning could be repeated just after the first acquisition due to the non-ionizing nature of REMS technique.

Other limitations of this study involve the following aspects: the population entirely composed of Caucasian women, the cross-sectional nature of the study itself and the lack of significant data on longitudinal evaluations of BMD_{US}.

Further studies showing a consistently high level of diagnostic accuracy in non-Caucasian populations and men are in fact needed, although there are no reasonable doubts to expect important variations as a function of sex and/or ethnic group. In the present study, which involved only Italian centers, male and non-Caucasian patients were excluded simply because their rate of DXA examinations in Italy is particularly low and the achievement of statistically significant numbers would have required an important lengthening of the study duration. However, international multicenter studies have already been started to overcome this issue. An analogous issue involves the fact that only 11% of the enrolled patients were obese (with 34% being overweight and 55% being normal- or under-weight). This BMI distribution reflects the typical characteristics of the populations of patients referred for a densitometric investigation, since low BMI is a recognized risk factor for osteoporosis and bone fragility fractures. In principle, increased BMI values could represent a challenge because of the augmented soft tissue thickness and further studies dedicated to obese patients are required for a detailed assessment. Anyway, REMS accuracy is actually not expected to be influenced by patient BMI as long as the patient morphology allows the proper setting of transducer focus and scan depth and the correct execution of subsequent scanning operations, which is typically feasible for all the patients with BMI up to 40 kg/m^2 (as verified also on the obese patients enrolled for this study).

The present study was a cross-sectional investigation aimed at comparing REMS results to DXA, assumed as a gold standard reference. Actually, prospective studies are needed in order to properly assess the association between REMS measurements and incident fractures, such to define whether the observed discrepancies between REMS and DXA can be finally labeled as REMS inaccuracies or they represent in some way a better fracture risk prediction, since BMD_{US} can be also slightly affected by bone quality characteristics that are not included in DXA-measured BMD. This will be one of the aims of future scheduled studies, in which we will also study the effectiveness of further novel dedicated REMS parameters, specifically thought to assess bone health and to estimate the possible consequent fracture risk on the basis of bone quality status, independently of BMD.

Longitudinal studies will be important also to verify whether the range of temporal responsiveness of BMD_{US} in a given patient is analogous to the corresponding variation range of DXA-measured BMD. While the slope of DXA vs REMS is close to unity, which would indicate that percentage values for

DXA precision errors can be compared to percentage errors for REMS, such longitudinal data on responsiveness will ultimately be required to confirm the smaller value of precision errors for REMS compared to DXA reported here.

Conclusion

Precision, repeatability, and the diagnostic accuracy of REMS investigations were assessed in comparison with DXA outcomes, in a multicenter clinical context, taking also into account possible errors in the performed DXA and REMS scans.

Obtained results showed that, when both DXA and REMS investigations were carried out in the strictest compliance with the corresponding guidelines and recommendations, REMS-measured BMD_{US} values resulted in good agreement with the corresponding DXA-measured BMD values for each considered anatomical site. This was also coupled with promising results in terms of measurement precision and repeatability.

Ongoing and future studies will include the evaluation of REMS effectiveness in male and non-Caucasian populations and the direct assessment of the degree of correlation between REMS outcomes and incident fractures, also exploiting the development of new parameters specifically dedicated to the assessment of bone structure quality.

Acknowledgments The authors thank Carla Signorini for her help and assistance in performing the DXA scans.

Funding information This work was partially funded by the National Research Council, Institute of Clinical Physiology, Lecce (Italy) within the project "Non-ionizing diagnoses in rheumatology—ECO MOC."

Compliance with ethical standards

Conflicts of interest Marco Di Paola, Davide Gatti, Ombretta Viapiana, Luisella Cianferotti, Loredana Cavalli, Carla Caffarelli, Eugenio Quarta, Paola Pisani, Giuseppe Girasole, Andrea Giusti, Monica Manfredini, Giovanni Arioli, Marco Matucci Cerinic, Gerolamo Bianchi, Ranuccio Nuti, Stefano Gonnelli, Maria Luisa Brandi, Maurizio Muratore, and Maurizio Rossini have no conflicts of interests.

Francesco Conversano owns stocks of Echolight Spa.

References

1. Kanis JA (2007) WHO Technical Report. University of Sheffield, UK, p 66
2. Johnell O, Kanis JA (2006) An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. Osteoporos Int 17:1726–1733
3. Cosman F, de Beur SJ, LeBoff MS, Lewiecki EM, Tanner B, Randall S, Lindsay R (2014) Clinician's guide to prevention and treatment of osteoporosis. Osteoporos Int 25:2359–2381
4. Lewiecki EM, Laster AJ (2006) Clinical review: clinical applications of vertebral fracture assessment by dual-energy X-ray absorptiometry. J Clin Endocrinol Metab 91(11):4215–4222
5. Abrahamsen B, van Staa T, Ariely R, Olson M, Cooper C (2009) Excess mortality following hip fracture: a systematic epidemiological review. Osteoporos Int 20(10):1633–1650
6. Office of the Surgeon General (US) (2004) Bone health and osteoporosis: a report of the Surgeon General. Office of the Surgeon General (US), Rockville (MD). Available from: http://www.ncbi.nlm.nih.gov/books/NBK45513/. Accessed December 2017
7. Cosman F, de Beur SJ, LeBoff MS, Lewiecki EM, Tanner B, Randall S, Lindsay R (2015) Erratum to: clinician's guide to prevention and treatment of osteoporosis. Osteoporos Int 26:2045–2047
8. International Osteoporosis Foundation (2000) How fragile is her future? Survey and Report. Available from:https://www.iofbonehealth.org/how-fragile-her-future. Accessed 12 June 2000
9. International Osteoporosis Foundation (2001) Osteoporosis in the European community: a call to action. Avalible from: https://www.iofbonehealth.org/osteoporosis-european-community-call-action. Accessed 18 December 2001
10. Quiros Roldan E, Brianese N, Raffetti E, Focà E, Pezzoli MC, Bonito A, Ferraresi A, Lanza P, Porcelli T, Castelli F (2017) Comparison between the gold standard DXA with calcaneal quantitative ultrasound based-strategy (QUS) to detect osteoporosis in an HIV infected cohort. Braz J Infect Dis 21:581–586
11. Schnitzer TJ, Wysocki N, Barkema D, Griffith J, Lent V, Romba M, Welbel R, Bhuva S, Manyam B, Linn S (2012) Calcaneal quantitative ultrasound compared with hip and femoral neck dual-energy X-ray absorptiometry in people with a spinal cord injury. PM R 4:748–755
12. Shin M-H, Kweon S-S, Park K-S, Heo H, Kim S-J, Nam H-S, Jeong S-K, Chung E-K, Choi J-S (2005) Quantitative ultrasound of the calcaneus in a Korean population: reference data and relationship to bone mineral density determined by peripheral dual X-ray absorptiometry. J Korean Med Sci 20:1011–1016
13. Paggiosi MA, Barkmann R, Gluer CC, Roux C, Reid DM, Felsenberg D, Bradburn M, Eastell R (2012) A European multicenter comparison of quantitative ultrasound measurement variables: the OPUS study. Osteoporos Int 23:2815–2828
14. Kwok T, Khoo CC, Leung J, Kwok A, Qin L, Woo J, Leung PC (2012) Predictive values of calcaneal quantitative ultrasound and dual energy X-ray absorptiometry for non-vertebral fracture in older men: results from the MrOS study (Hong Kong). Osteoporos Int 23:1001–1006
15. Liu JM, Ma LY, Bi YF, Xu Y, Huang Y, Xu M, Zhao HY, Sun LH, Tao B, Li XY, Wang WQ, Ning G (2012) A population-based study examining calcaneus quantitative ultrasound and its optimal cut-points to discriminate osteoporotic fractures among 9352 Chinese women and men. J Clin Endorinol Metab 97:800–809
16. Moayyeri A, Adams JE, Adler RA, Krieg MA, Hans D, Compston J, Lewiecki EM (2012) Quantitative ultrasound of the heel and fracture risk assessment: an updated meta-analysis. Osteoporos Int 23:143–153
17. Jiang YQ, Liu CC, Li RY, Wang WP, Ding H, Qi Q, Ta D, Dong J, Wang WQ (2014) Analysis of apparent integrated backscatter coefficient and backscatter spectral centroid shift in calcaneus in vivo for the ultrasonic evaluation of osteoporosis. Ultrasound Med Biol 40:1307–1317
18. Garra BS, Locher M, Felker S, Wear KA (2009) Measurements of ultrasonic backscattered spectral centroid shift from spine in vivo: methodology and preliminary results. Ultrasound Med Biol 35:165–168
19. Karjalainen JP, Riekkinen O, Toyras J, Hakulinen M, Kroger H, Rikkonen T, Salovaara K, Jurvelin JS (2012) Multi-site bone ultrasound measurements in elderly women with and without previous hip fractures. Osteoporos Int 23:1287–1295
20. Barkmann R, Dencks S, Laugier P, Padilla F, Brixen K, Ryg J, Seekamp A, Mahlke L, Bremer A, Heller M, Gluer CC (2010) Femur ultrasound (FemUS)—first clinical results on hip fracture discrimination and estimation of femoral BMD. Osteoporos Int 21:969–976

21. Schousboe JT, Riekkinen O, Karjalainen J (2017) Prediction of hip osteoporosis by DXA using a novel pulse-echo ultrasound device. Osteoporos Int 28:85–93
22. Karjalainen JP, Riekkinen O, Töyräs J, Jurvelin JS, Kröger H (2016) New method for point-of-care osteoporosis screening and diagnostics. Osteoporos Int 27:971–977
23. McCloskey EV, Kanis JA, Odén A, Harvey NC, Bauer D, González-Macias J, Hans D, Kaptoge S, Krieg MA, Kwok T, Marin F, Moayyeri A, Orwoll E, Gluër C, Johansson H (2015) Predictive ability of heel quantitative ultrasound for incident fractures: an individual-level meta-analysis. Osteoporos Int 26:1979–1987
24. Chan MY, Nguyen ND, Center JR, Eisman JA, Nguyen TV (2013) Quantitative ultrasound and fracture risk prediction in non-osteoporotic men and women as defined by WHO criteria. Osteoporos Int 24:1015–1022
25. Conversano F, Franchini R, Greco A, Soloperto G, Chiriacò F, Casciaro E, Aventaggiato M, Renna MD, Pisani P, Di Paola M, Grimaldi A, Quarta L, Quarta E, Muratore M, Laugier P, Casciaro S (2015) A novel ultrasound methodology for estimating spine mineral density. Ultrasound Med Biol 41:281–300
26. Casciaro S, Peccarisi M, Pisani P, Franchini R, Greco A, De Marco T, Grimaldi A, Quarta L, Quarta E, Muratore M, Conversano F (2016) An advanced quantitative echosound methodology for femoral neck densitometry. Ultrasound Med Biol 42:1337–1356
27. Casciaro S, Conversano F, Pisani P, Muratore M (2015) New perspectives in echographic diagnosis of osteoporosis on hip and spine. Clin Cases Miner Bone Metab 12:143–151
28. Messina C, Bandirali M, Sconfienza LM, D'Alonzo NK, Di Leo G, Papini GDE, Ulivieri FM, Sardanelli F (2015) Prevalence and type of errors in dual-energy x-ray absorptiometry. Eur Radiol 25:1504–1511
29. Engelke K, Gluer CC (2006) Quality and performance measures in bone densitometry. Part 1: errors and diagnosis. Osteoporos Int 17: 1283–1292
30. Shepherd JA, Schousboe JT, Broy SB, Engelke K, Leslie WD (2015) Executive summary of the 2015 ISCD position development conference on advanced measures from DXA and QCT: fracture prediction beyond BMD. J Clin Densitom 18:274–286
31. Altman DG, Bland JM (1983) Measurements in medicine: the analysis of method comparison studies. Statistician 32:307–317
32. Hui SL, Gao S, Zhou X-H, Johnston CC Jr, Lu Y, Glüer CC, Grampp S, Genant H (1997) Universal standardization of bone density measurements: a method with optimal properties for calibration among several instruments. J Bone Miner Res 12:1463–1470
33. Lu Y, Fuerst T, Hui S, Genant HK (2001) Standardization of bone mineral density at femoral neck, trochanter and Ward's triangle. Osteoporos Int 12:438–444
34. Hans DB, Shepherd JA, Schwartz EN, Reid DM, Blake GM, Fordham JN, Fuerst T, Hadji P, Itabashi A, Krieg MA, Lewiecki EM (2008) Peripheral dual-energy x-ray absorptiometry in the management of osteoporosis: the 2007 ISCD official positions. J Clin Densitom 11:188–206
35. Blake GM, Chinn DJ, Steel SA, Patel R, Panayiotou E, Thorpe J, Fordham JN, National Osteoporosis Society Bone Densitometry (2005) A list of device-specific thresholds for the clinical interpretation of peripheral x-ray absorptiometry examinations. Osteoporos Int 16:2149–2156
36. Hopkins SJ, Welsman JR, Knapp KM (2014) Short-term precision error in dual energy x-ray absorptiometry, bone mineral density and trabecular bone score measurements; and effects of obesity on precision error. Journal of Biomedical Graphics and Computing 4:8–14
37. Ravaud P, Reny JL, Giraudeau B, Porcher R, Dougados M, Roux C (1999) Individual smallest detectable difference in bone mineral density measurements. J Bone Miner Res 14:1449–1456
38. Trevisan C, Gandolini GG, Sibilla P, Penotti M, Caraceni MP, Ortolani S (1992) Bone mass measurement by DXA: influence of analysis procedures and interunit variation. J Bone Miner Res 7: 1373–1382
39. Larnach TA, Boyd SJ, Smart RC, Butler SP, Rohl PG, Diamond TH (1992) Reproducibility of lateral spine scans using dual energy X-ray absorptiometry. Calcif Tissue Int 51:255–258
40. Raffan E, Holden SL, Cullingham F, Hackett RM, Rawlings JM, German AJ (2006) Standardized positioning is essential for precise determination of body composition using dual-energy X-ray absorptiometry in dogs. J Nutr 136:1976S–1978S